Parametric Constructive Kripke-Semantics for Standard Multi-Agent Belief and Knowledge (Knowledge As Unbiased Belief)

Simon Kramer*
University of Luxembourg
simon.kramer@a3.epfl.ch

Joshua Sack
California State University Long Beach
joshua.sack@gmail.com

September 11, 2012

Abstract

We propose *parametric constructive Kripke-semantics* for multi-agent KD45-belief and S5-knowledge in terms of elementary set-theoretic constructions of two basic functional building blocks, namely *bias* (or *viewpoint*) and *visibility,* functioning also as the parameters of the doxastic and epistemic accessibility relation. The doxastic accessibility relates two possible worlds whenever the application of the composition of bias with visibility to the first world is equal to the application of visibility to the second world. The epistemic accessibility is the transitive closure of the union of our doxastic accessibility and its converse. Therefrom, accessibility relations for *common* and *distributed* belief and knowledge can be constructed in a standard way. As a result, we obtain *a general definition of knowledge in terms of belief* that enables us to view S5-knowledge as accurate (unbiased and thus true) KD45-belief, negation-complete belief and knowledge as exact KD45-belief and S5-knowledge, respectively, and perfect S5-knowledge as precise (exact and accurate) KD45-belief, and all this *generically* for arbitrary functions of bias and visibility. Our results can be seen as a semantic complement to previous foundational results by Halpern et al. about the (un)definability and (non-)reducibility of knowledge in terms of and to belief, respectively.

Keywords: parametric constructive Kripke-semantics, doxastic & epistemic logic, knowledge as a form of belief, multi-agent distributed systems, relational semantics of modal logic.

*funded with Grant AFR 894328 from the National Research Fund Luxembourg cofunded under the Marie-Curie Actions of the European Commission (FP7-COFUND)

1 Introduction

In [HSS09], the problem of defining knowledge in terms of belief is studied from a modal logic perspective, where the authors show that "if knowledge satisfies any set of axioms contained in S5, then it cannot be explicitly defined in terms of belief. S5 knowledge can be implicitly defined by belief, but not reduced to it." Thereby, the standard notions of explicit and implicit definability from first-order logic are "lifted to the definability of modalities in modal logics in a straightforward way," so that "explicit definability is equivalent to the combination of implicit definability and reducibility." More precisely, [HSS09]:

> Consider a logic Λ for knowledge and belief. Knowledge is explicitly defined in Λ if there is a formula DK (for "definition of knowledge") in Λ of the form $Kp \leftrightarrow \delta$, where δ is a formula that does not mention the knowledge operator. Knowledge is implicitly defined in Λ if, roughly speaking, Λ "determines" knowledge uniquely. Syntactically, this determination means that any two modal operators for knowledge that satisfy Λ must be equivalent. Semantically, this means that two Kripke models of Λ with the same set of worlds that agree on the interpretation of belief (and on the interpretations of all primitive propositions) must agree also on the interpretation of knowledge.

Our contribution is to make the definability of S5-knowledge in terms of KD45-belief *function-parametric* as well as *semantic-constructive* (cf. Definition 3–8). More precisely, we propose function-parametric constructive Kripke-semantics for multi-agent KD45-belief and S5-knowledge in terms of elementary set-theoretic constructions of two basic functional building blocks, namely (cf. Definition 2):

- *bias,* or viewpoint translocation (necessarily idempotent, e.g., the constant functions), and
- *visibility* transformation, for example:
  - point confounding (non-injective when non-trivial), and/or
  - point confusing (or permuting, bijective on a sub-domain)

functioning also as the parameters of the doxastic and epistemic accessibility relation. Note that we mean "*set-theoretically* constructive," not "intuitionistic," in loose analogy with the set-theoretically constructive rather than the purely axiomatic definition of numbers [Fef89] or ordered pairs.[1] That is:

- our *epistemic* accessibility is the transitive closure of the union of our doxastic accessibility and its converse (cf. Definition 4);
- our *doxastic* accessibility relates two possible worlds whenever the application of the composition of bias with visibility to the first world is equal to the application of visibility to the second world (cf. Definition 3).

[1] E.g., the now standard definition by Kuratowski or other well-known definitions [Mos06].

As a result, our constructions enable us to view

- S5-knowledge as accurate (unbiased and thus true) KD45-belief,
- *negation-complete* belief and knowledge as exact KD45-belief and S5-knowledge, respectively,
- *perfect* S5-knowledge as precise (exact and accurate) KD45-belief, and

all this *generically* for arbitrary functions of bias and visibility in our sense (cf. Theorem 3). In comparison, recall from [FHMV95] the by-now classic constructive definition of agent-centric (say in agent a) epistemic accessibility $\equiv_a$ as state (say s and s') indistinguishability $s \equiv_a s'$ defined in terms of the equality between the projection $\pi_a(s)$ of s onto a's view and the projection $\pi_a(s')$:

Definition 1 (Epistemic accessibility as state indistinguishability [FHMV95])**.**

$$s \equiv_a s' \text{ by definition, if and only if } \pi_a(s) = \pi_a(s').$$

Thus $\equiv_a$ is defined to be the kernel of π_a [SD08]. This definition is constructive in the sense that it not merely abstractly stipulates $\equiv_a$ to be an equivalence relation (that would be the standard modal-logical methodology [MV07]) but it actually concretely constructs $\equiv_a$ in terms of the set-theoretic building block π_a, a projection function (state visibility as state projection), which *forces* $\equiv_a$ to be an equivalence relation. (For more examples of more complex, constructive definitions of agent-centric accessibility relations, see [Kra12a, Kra12c, Kra12b].) It is not at all obvious how to recover a definition of doxastic accessibility from this indistinguishability definition of epistemic accessibility. Nevertheless we present a simple, generic, and thus general solution to this important problem. Our solution is general in the sense that the extent to which it can be applied is the entire semantic scope (models) of standard doxastic and epistemic logic (cf. [MV07] and [HR10] for overviews), thanks to our soundness and completeness results in the sense of Theorem 1 and 2. Moreover, our proofs for the solution are simple, which increases its value. Here, the difficulty was to *find* our general definition of knowledge in terms of belief, which has even the feature of being generic thanks to its function parameters. Our findings can be seen as a semantic complement to previous foundational results by Halpern et al. about the (un)definability and (non-)reducibility of knowledge in terms of and to belief, respectively.

2 Parametric constructions and results

Let

- $\mathcal{S}$ designate a set of system states in computer-science, points in modal-logical, or possible worlds in philosophical terminology;
- $\mathrm{id}_{\mathcal{S}'} := \{(s, s) \mid s \in \mathcal{S}'\}$ the identity function on $\mathcal{S}' \subseteq \mathcal{S}$;

- $\mathrm{Im}(R) := \{s' \in \mathcal{S} \mid \text{there is } s \in \mathcal{S} \text{ such that } s\, R\, s'\}$ the image of some (possibly functional) relation $R \subseteq \mathcal{S} \times \mathcal{S}$.

Further let ":iff" abbreviate "by definition, if and only if".

Definition 2 (Doxastic-epistemic function pair)**.** *Two functions $f : \mathcal{S} \to \mathcal{S}$ and $g : \mathrm{Im}(f) \to \mathrm{Im}(f)$ form a* doxastic-epistemic function pair (f, g) on $\mathcal{S}$ *:iff*

- *for all $s, s' \in \mathcal{S}$, if $g(f(s)) = f(s')$ then $g(f(s')) = f(s')$;*
- ***or, equivalently,** g is idempotent, i.e., $g \circ g = g$.*

Fact 1. *For all doxastic-epistemic function pairs (f, g) on $\mathcal{S}$ and $s \in \mathcal{S}$ there is $s' \in \mathcal{S}$ such that $g(f(s)) = f(s')$.*

Proof. By the definitional fact that f is a totally defined operation on $\mathcal{S}$ and g is a totally defined operation on $\mathrm{Im}(f)$; $(\mathrm{Im}(g) \subseteq \mathrm{Im}(f))$. □

We shall use the two constraints in Definition 2 interchangeably; the two constraints are indeed equivalent, as asserts the following proposition.

Proposition 1. *The two alternative constraints in Definition 2 are equivalent.*

Proof. Let $f : \mathcal{S} \to \mathcal{S}$ and $g : \mathrm{Im}(f) \to \mathrm{Im}(f)$. For the if-direction, suppose that for all $s, s' \in \mathcal{S}$, if $g(f(s)) = f(s')$ then $g(f(s')) = f(s')$. Further let $s \in \mathcal{S}$. By Fact 1, there is $s' \in \mathcal{S}$ such that $g(f(s)) = f(s')$. Hence $g(f(s)) = g(f(s'))$, and also $g(g(f(s))) = g(f(s'))$. Hence $g(g(f(s))) = g(f(s))$. For the only-if-direction, suppose that $g \circ g = g$, and let $s, s' \in \mathcal{S}$. Further suppose that $g(f(s)) = f(s')$. Hence $g(g(f(s))) = g(f(s'))$. Hence $g(f(s)) = g(f(s'))$ by the idempotency of g. Hence $g(f(s')) = f(s')$ by the last supposition. □

Definition 3 (Function-Parametric *Doxastic* Accessibility)**.** *Let (f, g) designate a doxastic-epistemic function pair on $\mathcal{S}$. Then we define our (f, g)*-parametric **doxastic** accessibility relation $\mathrm{D}_f^g \subseteq \mathcal{S} \times \mathcal{S}$ *such that for all $s, s' \in \mathcal{S}$,*

$$\boxed{s\, \mathrm{D}_f^g\, s' \;\text{:iff}\; g(f(s)) = f(s').}$$

The following main *adequacy theorem* asserts first that for all doxastic-epistemic function pairs (f, g), D_f^g is indeed a standard doxastic accessibility relation, and second that for all standard doxastic accessibility relations R, a doxastic-epistemic function pair (f, g) can be constructed such that $R = \mathrm{D}_f^g$.

Theorem 1 (The **KD45** Accessibility Schema)**.**

1. ***Soundness:*** *If (f, g) is a doxastic-epistemic function pair on $\mathcal{S}$ then for all $s \in \mathcal{S}$:*
   (a) *there is $s' \in \mathcal{S}$ such that $s\, \mathrm{D}_f^g\, s'$ (Seriality/Totality)*
   (b) *for all $s', s'' \in \mathcal{S}$:*
      i. *if $s\, \mathrm{D}_f^g\, s'$ and $s'\, \mathrm{D}_f^g\, s''$ then $s\, \mathrm{D}_f^g\, s''$ (Transitivity)*

ii. if $s \ \mathrm{D}_f^g \ s'$ and $s \ \mathrm{D}_f^g \ s''$ then $s' \ \mathrm{D}_f^g \ s''$ (Euclideanness)

2. ***Completeness:*** *If $\emptyset \neq R \subseteq \mathcal{S}\times\mathcal{S}$ is serial, transitive, and Euclidean then there is a doxastic-epistemic function pair (f,g) on $\mathcal{S}$ such that (f,g) is constructible from R and $R = \mathrm{D}_f^g$.*

Proof. For soundness, assume that (f,g) is a doxastic-epistemic function pair on $\mathcal{S}$. Then 1.a holds by Fact 1. For 1.b, let $s, s', s'' \in \mathcal{S}$ and suppose that $g(f(s)) = f(s')$. Hence $g(f(s')) = f(s')$ by the first alternative definitional constraint on f and g. For 1.b.i, further derive that $g(f(s)) = g(f(s'))$ by the last supposition, and then further suppose that $g(f(s')) = f(s'')$. Hence $g(f(s)) = f(s'')$. For 1.b.ii, suppose that $g(f(s)) = f(s'')$. Consequently, $f(s') = f(s'')$ by the last supposition and the first supposition of 1.b, and then $g(f(s')) = f(s'')$ by the very first derivation.

For completeness, let $\emptyset \neq R \subseteq \mathcal{S}\times\mathcal{S}$ and suppose that R is serial, transitive, and Euclidean. Then $\equiv := \{(s,s') \in R \mid s \in \mathrm{Im}(R)\} \subseteq R$ is an equivalence relation (i.e., a relation that is reflexive, transitive, and Euclidean) on $\mathrm{Im}(R)$:

- $\equiv$ is reflexive: Let $s \in \mathrm{Im}(R)$, i.e., there is $s' \in \mathcal{S}$ such that $s' \ R \ s$. Thus $s \in \mathcal{S}$. Hence there is $s'' \in \mathcal{S}$ such that $s \ R \ s''$ by the seriality of R. Hence $s' \ R \ s''$ by the transitivity of R. Hence $s'' \ R \ s$ by the Euclideanness of R. Hence $s \ R \ s$ by the transitivity of R. And since $s \in \mathrm{Im}(R)$, $s \equiv s$.

- $\equiv$ is transitive and Euclidean by inheritance, i.e., simply because R is.

For each $s \in \mathrm{Im}(R)$, choose $c_s \in [s]_\equiv$ such that for all $s' \in \mathcal{S}$, if $[s]_\equiv = [s']_\equiv$ then $c_s = c_{s'}$. Observe that for all $s', s'' \in \mathcal{S}$, if $s \ R \ s'$ and $s \ R \ s''$ then there is $C \in \mathrm{Im}(R)/_\equiv$ such that $s', s'' \in C$. This is because $s' \ R \ s''$ by the Euclideanness of R and because $s', s'' \in \mathrm{Im}(R)$ (thus $s' \equiv s''$). Now define two functions $f : \mathcal{S} \to \mathcal{S}$ and $g : \mathrm{Im}(f) \to \mathrm{Im}(f)$ such that:

$$f : s \mapsto \begin{cases} c_s & \text{if } s \in \mathrm{Im}(R)\,, \\ s & \text{if } s \notin \mathrm{Im}(R)\,; \end{cases}$$

$$g : s \mapsto \begin{cases} s & \text{if } s \in \mathrm{Im}(R)\,, \\ c_{s'} & \text{if } s \notin \mathrm{Im}(R) \text{ and } s \ R \ s'. \end{cases}$$

Notice that g is well defined, i.e., it does not matter which s' we choose, since for all $s'' \in \mathcal{S}$, if $s \ R \ s'$ and $s \ R \ s''$ then there is $C \in \mathrm{Im}(R)/_\equiv$ such that $s', s'' \in C$.

We will now see that $R = \mathrm{D}_f^g$. So let $s, s' \in \mathcal{S}$.

- Suppose that $s \ R \ s'$. Thus $s' \in \mathrm{Im}(R)$. Hence $f(s') = c_{s'}$.
  - Suppose that $s \in \mathrm{Im}(R)$. Hence $s \equiv s'$ by the first two suppositions (thus $c_s = c_{s'}$), and $g(f(s)) = g(c_s) = c_s$. Hence $g(f(s)) = f(s')$.
  - Now suppose that $s \notin \mathrm{Im}(R)$. Hence $g(f(s)) = g(s) = c_{s''}$ for some $s'' \in \mathcal{S}$ such that $s \ R \ s''$. Hence $s' \ R \ s''$ by the Euclideanness of R, and then $s' \equiv s''$. Thus $c_{s'} = c_{s''}$. Hence $g(f(s)) = f(s')$.

- Conversely suppose that $g(f(s)) = f(s')$. Notice in the definition of f and g that $\mathrm{Im}(g) \subseteq \mathrm{Im}(R)$ and that $s' \in \mathrm{Im}(R)$ if and only if $f(s') \in \mathrm{Im}(R)$. Hence $s' \in \mathrm{Im}(R)$, and thus $f(s') = c_{s'}$. Hence $g(f(s)) = c_{s'}$.
  - Suppose that $s \in \mathrm{Im}(R)$. Hence $g(f(s)) = g(c_s) = c_s$. Hence $c_s = c_{s'}$ and thus $s \equiv s'$. Hence $s \mathrel{R} s'$.
  - Now suppose that $s \notin \mathrm{Im}(R)$. Hence $g(f(s)) = g(s) = c_{s''}$ for some $s'' \in \mathcal{S}$ such that $s \mathrel{R} s''$. Hence $c_{s''} = c_{s'}$, and thus $s'' \equiv s'$. Hence $s'' \mathrel{R} s'$. Hence $s \mathrel{R} s'$ by the transitivity of R.

□

The following proposition gives a functional characterisation of (i.e., a necessary and sufficient equational condition for) the symmetry (and hence the property of being an equivalence relation) of (f, g)-parametric doxastic accessibility relations. (Seriality, symmetry, and transitivity jointly imply reflexivity.)

Proposition 2 (Doxastic symmetry characterisation)**.** *For all (f, g)-parametric doxastic accessibility relations $\mathrm{D}_f^g \subseteq \mathcal{S} \times \mathcal{S}$,*

$$\mathrm{D}_f^g = (\mathrm{D}_f^g)^{-1} \text{ if and only if } g = \mathrm{id}_{\mathrm{Im}(f)} \,.$$

Proof. The if-direction is immediate. For the only-if-direction, suppose that $\mathrm{D}_f^g = (\mathrm{D}_f^g)^{-1}$, i.e., for all $s, s' \in \mathcal{S}$, $g(f(s)) = f(s')$ if and only if $g(f(s')) = f(s)$. Further, let $s \in \mathcal{S}$. Hence there is $s' \in \mathcal{S}$ such that $g(f(s)) = f(s')$ by Fact 1. Hence $g(f(s')) = f(s)$ by the first supposition, and also $g(g(f(s))) = g(f(s'))$. Hence, $g(f(s)) = g(f(s'))$ by the idempotency of g, and then $g(f(s)) = f(s)$. □

The following is our general definition of knowledge in terms of belief.

Definition 4 (Function-Parametric *Epistemic* Accessibility)**.** *Let D_f^g designate an (f, g)-parametric* doxastic *accessibility. Then we define our (f, g)-*parametric **epistemic** accessibility relation $\mathrm{E}_f^g \subseteq \mathcal{S} \times \mathcal{S}$ *such that*

$$\boxed{\mathrm{E}_f^g := (\mathrm{D}_f^g \cup (\mathrm{D}_f^g)^{-1})^+,}$$

where '${}^{-1}$' designates the converse and '${}^{+}$' the transitive-closure operation.

The following *adequacy theorem* asserts first that for all doxastic-epistemic function pairs (f, g), E_f^g is indeed a standard epistemic accessibility relation, and second that for all standard epistemic accessibility relations $\equiv$, a doxastic-epistemic function pair (f, g) can be constructed such that $\equiv\, = \mathrm{E}_f^g$.

Theorem 2 (The **S5** Accessibility Schema)**.**

1. ***Soundness:*** *If D_f^g is an (f, g)-parametric accessibility relation then E_f^g is the smallest equivalence relation containing D_f^g.*

2. ***Completeness:*** *If* $\emptyset \neq \equiv \subseteq \mathcal{S}\times\mathcal{S}$ *is an equivalence relation then there is a doxastic-epistemic function pair* (f, g) *on* $\mathcal{S}$ *such that* (f, g) *is constructible from* $\equiv$ *and* $\equiv = \mathrm{E}_f^g$.

Proof. For soundness, consider that since D_f^g is serial and *transitive* by construction, and since the converse operation preserves the seriality and transitivity of D_f^g, E_f^g is so too. Finally, since E_f^g is *symmetric* by construction, E_f^g is also *reflexive,* and thus an equivalence relation containing D_f^g. (Seriality, symmetry, and transitivity jointly imply reflexivity.) To see that E_f^g is the *smallest* such relation, recall from [SD08] that for arbitrary $R \subseteq \mathcal{S} \times \mathcal{S}$, the relation $(R \cup (R)^{-1} \cup \mathrm{id}_{\mathcal{S}})^*$ is the smallest equivalence relation containing R, where '$*$' is the reflexive-transitive-closure operation. However, we can spare $\mathrm{id}_{\mathcal{S}}$ and the reflexive closure, since $\mathrm{id}_{\mathcal{S}} \subseteq \mathrm{E}_f^g$ is reflexive by construction.

For completeness, suppose that $\emptyset \neq \equiv \subseteq \mathcal{S} \times \mathcal{S}$ is an equivalence relation. Then for each equivalence class $C \in \mathcal{S}/_{\equiv}$, choose $c \in C$ and define $\pi(s) := c$ for all $s \in C$. Clearly, $\equiv = \mathrm{D}_\pi^{\mathrm{id}_{\mathrm{Im}(\pi)}}$, and $\mathrm{D}_\pi^{\mathrm{id}_{\mathrm{Im}(\pi)}} = \mathrm{E}_\pi^{\mathrm{id}_{\mathrm{Im}(\pi)}}$ (see also Proposition 3.2). Thus $\equiv = \mathrm{E}_\pi^{\mathrm{id}_{\mathrm{Im}(\pi)}}$. □

The following proposition is the basis for our third main result, namely Theorem 3.

Proposition 3 (Doxastic-epistemic accessibility inclusions)**.** *For all doxastic-epistemic function pairs* (f, g) *on* $\mathcal{S}$:

1. $\mathrm{D}_f^g \subseteq \mathrm{E}_f^g$;
2. $\mathrm{D}_f^g = \mathrm{E}_f^g$ *if and only if* $g = \mathrm{id}_{\mathrm{Im}(f)}$;
3. $\mathrm{id}_{\mathcal{S}} = \mathrm{E}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{id}_{\mathcal{S}}} = \mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{id}_{\mathcal{S}}}$.

Proof. For 1, inspect definitions. 2 follows from Proposition 2 and the definition of E_f^g . For 3, inspect 2. □

Definition 5 (Doxastic-epistemic similarity type)**.** *Let*

- $\mathcal{P} \neq \emptyset$ *designate some set of* atomic propositions P ;
- $\mathcal{T}$ *a set of* types T *such that* $\mathtt{S} \in \mathcal{T}$;
- $\mathcal{G}$ *a set of typed function names* $\boldsymbol{g} : T \to T'$ *(abbreviated as* $\boldsymbol{g}$ *when clear from context) such that* $(\mathtt{id}_T : T \to T) \in \mathcal{G}$ *for all* $T \in \mathcal{T}$;
- $\mathcal{F}$ *a set of typed function names* $\boldsymbol{f} : \mathtt{S} \to T$ *(abbreviated as* $\boldsymbol{f}$ *when clear from context) such that* $(\mathtt{id}_\mathtt{S} : \mathtt{S} \to \mathtt{S}) \in \mathcal{F}$ *and if* $(\boldsymbol{f} : \mathtt{S} \to T) \in \mathcal{F}$ *and* $(\boldsymbol{h} : T \to T') \in \mathcal{F} \cup \mathcal{G}$ *then* $(\boldsymbol{h} \circ_T \boldsymbol{f} : \mathtt{S} \to T') \in \mathcal{F}$;
- $\Pi_B, \Pi_K \subseteq \Pi := \{(\boldsymbol{f} : T \to T', \boldsymbol{g} : T' \to T'') \in \mathcal{F} \times \mathcal{G}\}$ belief-label *and* knowledge-label sets, *and the so-called* label set*, respectively.*

Then,

$$\Sigma := (\mathcal{P}, \mathcal{T}, \mathcal{G}, \mathcal{F}, \Pi_B, \Pi_K)$$

is a doxastic-epistemic similarity type*.*

Note the above-introduced notational conventions: we use f, g, and h as meta-variables for typed function names, and f, g, and h as meta-variables for typed functions; $\mathtt{id}_{\mathtt{S}}$ is an example of a (typed) function name, and $\mathrm{id}_{\mathcal{S}}$ is an example of a (typed) function.

Definition 6 (Functional doxastic-epistemic language)**.** *Given a doxastic-epistemic similarity type Σ with set $\mathcal{P}$ of atomic propositions, and belief- and knowledge-label set Π_B and Π_K, respectively,*

$$\begin{array}{rcl} \mathcal{L}(\Sigma) \ni \phi & ::= & P\ (P \in \mathcal{P}) \mid \neg\phi \mid \phi \wedge \phi \\ & & \mid \mathsf{B}^{\mathsf{g}}_{\mathsf{f}}(\phi)\ ((\mathsf{f}, \mathsf{g}) \in \Pi_B) \\ & & \mid \mathsf{K}^{\mathsf{g}}_{\mathsf{f}}(\phi)\ ((\mathsf{f}, \mathsf{g}) \in \Pi_K) \end{array}$$

is the doxastic-epistemic language over Σ*.*

We intend for the operators $\mathsf{B}^{\mathsf{g}}_{\mathsf{f}}$ and $\mathsf{K}^{\mathsf{g}}_{\mathsf{f}}$ to have flexible readings, though they generally relate to belief and knowledge, respectively. We may associate every pair (f, g) with an agent a in some given set $\mathcal{A}$ of agents (as we do in Proposition 6). When doing so, we may read $\mathsf{B}^{\mathsf{g}}_{\mathsf{f}}$ as "agent a believes that ϕ" and $\mathsf{K}^{\mathsf{g}}_{\mathsf{f}}$ as "agent a knows that ϕ," where a is the agent associated with the pair (f, g).

Definition 7 (Functional doxastic-epistemic models)**.** *Given a doxastic-epistemic similarity type $\Sigma = (\mathcal{P}, \mathcal{T}, \mathcal{G}, \mathcal{F}, \Pi_B, \Pi_K)$, let*

- $\langle \mathcal{S}, \iota \rangle$ *be a so-called* Σ-instantiation structure on $\mathcal{S}$ *with an interpretation function ι for types $T \in \mathcal{T}$ and typed function names constrained such that:*
  - $\iota(\mathtt{S}) = \mathcal{S}$ *and* $\iota(T) \subseteq \mathcal{S}$;
  - $\boxed{\begin{array}{rcl} \iota(\mathtt{id}_T : T \to T) & := & \mathrm{id}_{\iota(T)}\,, \\ \iota(\mathsf{h} \circ_{T'} \mathsf{f} : T \to T'') & := & \iota(\mathsf{h} : T' \to T'') \circ \iota(\mathsf{f} : T \to T')\,; \end{array}}$
  - $\iota(\mathsf{h} : T \to T')$ *is a function* $h : \iota(T) \to \iota(T')$ *such that if* $\mathsf{h} \in \mathcal{G}$ *then* h *is idempotent;*
- $\langle \mathcal{S}, \{\mathrm{D}^{\iota(\mathsf{g})}_{\iota(\mathsf{f})}\}_{(\mathsf{f},\mathsf{g})\in\Pi_B}, \{\mathrm{E}^{\iota(\mathsf{g})}_{\iota(\mathsf{f})}\}_{(\mathsf{f},\mathsf{g})\in\Pi_K} \rangle$ *a* doxastic-epistemic Σ-frame on $\langle \mathcal{S}, \iota \rangle$;
- $\mathcal{V} : \mathcal{P} \to 2^{\mathcal{S}}$ *a standard modal* valuation function *[BvB07], mapping each atomic proposition P to the set of states where P is considered true.*

Then,

$$\mathfrak{S} := \langle \mathcal{S}, \{\mathrm{D}^{\iota(\mathsf{g})}_{\iota(\mathsf{f})}\}_{(\mathsf{f},\mathsf{g})\in\Pi_B}, \{\mathrm{E}^{\iota(\mathsf{g})}_{\iota(\mathsf{f})}\}_{(\mathsf{f},\mathsf{g})\in\Pi_K}, \mathcal{V} \rangle$$

is the **doxastic-epistemic Σ-model on $\langle \mathcal{S}, \iota \rangle$ and $\mathcal{V}$***, and* $(\mathfrak{S}, s)$ *a* **pointed** doxastic-epistemic Σ-model on $\langle \mathcal{S}, \iota \rangle$ and $\mathcal{V}$ *for any* $s \in \mathcal{S}$*.*

Table 1: Doxastic-epistemic satisfaction relation

$(\mathfrak{S}, s) \models P$	:iff	$s \in \mathcal{V}(P)$
$(\mathfrak{S}, s) \models \neg\phi$	:iff	not $(\mathfrak{S}, s) \models \phi$
$(\mathfrak{S}, s) \models \phi \wedge \phi'$	:iff	$(\mathfrak{S}, s) \models \phi$ and $(\mathfrak{S}, s) \models \phi'$
$(\mathfrak{S}, s) \models \mathsf{B}^{g}_{f}(\phi)$	:iff	for all $s' \in \mathcal{S}$, if $s \ \mathrm{D}^{\iota(g)}_{\iota(f)} \ s'$ then $(\mathfrak{S}, s') \models \phi$
$(\mathfrak{S}, s) \models \mathsf{K}^{g}_{f}(\phi)$	:iff	for all $s' \in \mathcal{S}$, if $s \ \mathrm{E}^{\iota(g)}_{\iota(f)} \ s'$ then $(\mathfrak{S}, s') \models \phi$

Definition 8 (Functional doxastic-epistemic logic)**.** *Given a doxastic-epistemic similarity type* $\Sigma = (\mathcal{P}, \mathcal{T}, \mathcal{G}, \mathcal{F}, \Pi_B, \Pi_K)$*, define*

- *a standard* satisfaction relation $\models$ *between pointed doxastic-epistemic* Σ*-models and their languages* $\mathcal{L}(\Sigma)$ *as in Table 1;*
- $\mathfrak{S} \models \phi$ *:iff for all* $s \in \mathcal{S}$*,* $(\mathfrak{S}, s) \models \phi$*;*
- $\models \phi$ *:iff for all doxastic-epistemic* Σ*-models* $\mathfrak{S}$*,* $\mathfrak{S} \models \phi$*.*

Proposition 4 (KD45-belief modality)**.** *For all* $(f, g) \in \Pi_B$*,* B^{g}_{f} *is a* KD45-belief modality. *That is:*

1. $\models \mathsf{B}^{g}_{f}(\phi \to \phi') \to (\mathsf{B}^{g}_{f}(\phi) \to \mathsf{B}^{g}_{f}(\phi'))$ *(Kripke's law, K)*
2. $\models \neg\mathsf{B}^{g}_{f}(\bot)$ *(equivalently,* $\models \mathsf{B}^{g}_{f}(\phi) \to \neg\mathsf{B}^{g}_{f}(\neg\phi)$*) (belief consistency, D)*
3. $\models \mathsf{B}^{g}_{f}(\phi) \to \mathsf{B}^{g}_{f}(\mathsf{B}^{g}_{f}(\phi))$ *(positive introspection, **4**)*
4. $\models \neg\mathsf{B}^{g}_{f}(\phi) \to \mathsf{B}^{g}_{f}(\neg\mathsf{B}^{g}_{f}(\phi))$ *(negative introspection, **5**)*
5. *if* $\models \phi$ *then* $\models \mathsf{B}^{g}_{f}(\phi)$ *(necessitation, N)*

Proof. By Theorem 1.1. K and N are forced by Kripke-semantics. The D-law corresponds to seriality, the 4-law to transitivity, and the 5-law to Euclideanness. □

Proposition 5 (S5-knowledge modality)**.** *For all* $(f, g) \in \Pi_K$*,* K^{g}_{f} *is an* S5-belief modality. *That is:*

1. $\models \mathsf{K}^{g}_{f}(\phi \to \phi') \to (\mathsf{K}^{g}_{f}(\phi) \to \mathsf{K}^{g}_{f}(\phi'))$ *(Kripke's law, K)*
2. $\models \mathsf{K}^{g}_{f}(\phi) \to \phi$ *(truth law, T)*
3. $\models \mathsf{K}^{g}_{f}(\phi) \to \mathsf{K}^{g}_{f}(\mathsf{K}^{g}_{f}(\phi))$ *(positive introspection, **4**)*
4. $\models \neg\mathsf{K}^{g}_{f}(\phi) \to \mathsf{K}^{g}_{f}(\neg\mathsf{K}^{g}_{f}(\phi))$ *(negative introspection, **5**)*
5. *if* $\models \phi$ *then* $\models \mathsf{K}^{g}_{f}(\phi)$ *(necessitation, N)*

Proof. By Theorem 2.1. The T-law corresponds to reflexivity. □

The following theorem summarises our main results.

Theorem 3 (Doxastic-epistemic modality conditionals)**.**

1. $\boxed{\models \mathsf{K}_{f}^{\mathtt{id}_{\mathtt{Im}(f)}}(\phi) \leftrightarrow \mathsf{B}_{f}^{\mathtt{id}_{\mathtt{Im}(f)}}(\phi)}$ *(knowledge as unbiased belief)*
2. $\models \mathsf{B}_{f}^{\mathtt{id}_{\mathtt{Im}(f)}}(\phi) \rightarrow \phi$ *(unbiased belief is true belief)*
3. $\models \phi \rightarrow \mathsf{K}_{\mathtt{id}_{\mathtt{S}}}^{\mathtt{id}_{\mathtt{S}}}(\phi)$ *(perfect knowledge)*
4. $\models \mathsf{K}_{\mathtt{id}_{\mathtt{S}}}^{\mathtt{id}_{\mathtt{S}}}(\phi) \leftrightarrow \mathsf{B}_{\mathtt{id}_{\mathtt{S}}}^{\mathtt{id}_{\mathtt{S}}}(\phi)$ *(perfect knowledge as precise belief)*
5. $\models \mathsf{K}_{f}^{g}(\phi) \rightarrow \mathsf{B}_{f}^{g}(\phi)$ *(knowledge implies belief, like in [HSS09])*
6. *bias cancellation:*
   (a) $\mathfrak{S} \models \mathsf{K}_{f}^{g}(\phi) \leftrightarrow \mathsf{B}_{f}^{g}(\phi)$ *if and only if* $\iota(g) = \mathrm{id}_{\mathrm{Im}(\iota(f))}$,
   (b) $\iota(g) = \mathrm{id}_{\mathrm{Im}(\iota(f))}$ *if and only if* $\iota(g)$ *is injective;*
7. $\boxed{\models \neg\mathsf{B}_{\mathtt{id}_{\mathtt{S}}}^{g}(\phi) \rightarrow \mathsf{B}_{\mathtt{id}_{\mathtt{S}}}^{g}(\neg\phi)}$ *(negation-complete belief)*
8. *negation-complete knowledge:*
   (a) $\mathfrak{S} \models \neg\mathsf{K}_{\mathtt{id}_{\mathtt{S}}}^{g}(\phi) \rightarrow \mathsf{K}_{\mathtt{id}_{\mathtt{S}}}^{g}(\neg\phi)$ *if and only if* $\iota(g)$ *is injective,*
   (b) *negation-complete knowledge coincides with perfect knowledge.*

Proof. 1 follows from Proposition 3.2; 2 from 1 and Proposition 5.2; 3 and 4 from the left and right equation in Proposition 3.3, respectively; (4 also as an instance of 1;) 5 from Proposition 3.1; 6.a from Proposition 3.2; 6.b from the fact that an idempotent function that is also injective must be the identity; 7 from the functionality of $\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\iota(g)}$ for any ι; 8.a from the functionality of $\mathrm{E}_{\mathrm{id}_{\mathcal{S}}}^{\iota(g)}$ for injective $\iota(g)$; and 8.b from 8.a, 6.b, and 3. $\square$

The following proposition establishes formal correspondences to related work.

Proposition 6 (Related work)**.**

1. Epistemic accessibility as state indistinguishability [FHMV95]: *Let*

$$\mathcal{S} \ni s ::= \mathsf{0} \mid \alpha_a(s)\,,$$

where $\mathsf{0}$ *designates a zero data point (e.g., an initial state) and* α_a *an action performed by agent* $a \in \mathcal{A}$ *for some finite set* $\mathcal{A} \neq \emptyset$ *of agents, and define the visibility* π_a *in Definition 1 on* $\mathcal{S}$ *as*

$$\pi_a(\mathsf{0}) := \mathsf{0}$$

$$\pi_a(\alpha_b(s)) := \begin{cases} \alpha_b(\pi_a(s)) & \text{if } a = b, \text{ and} \\ \pi_a(s) & \text{otherwise.} \end{cases}$$

Then,

$$\boxed{\mathrm{E}_{\pi_a}^{\mathrm{id}_{\mathrm{Im}(\pi_a)}} = \equiv_a .}$$

Thus we can reconstruct the standard agent-centric epistemic modality K_a *[FHMV95] in our framework with the following simple definition*

$$\mathsf{K}_a(\phi) := \mathsf{K}_{\mathtt{pi}_a}^{\mathtt{id}_{\mathtt{Im}(\mathtt{pi}_a)}}(\phi)$$

for doxastic-epistemic similarity types such that $\mathcal{T} := \{\mathtt{S}\} \cup \{\mathtt{Im}(\mathtt{pi}_a) \mid a \in \mathcal{A}\}$, $\mathcal{G} := \{\mathtt{id}_{\mathtt{Im}(\mathtt{pi}_a)} : \mathtt{Im}(\mathtt{pi}_a) \to \mathtt{Im}(\mathtt{pi}_a) \mid a \in \mathcal{A}\}$, $\mathcal{F} := \{\mathtt{pi}_a : \mathtt{S} \to \mathtt{Im}(\mathtt{pi}_a) \mid a \in \mathcal{A}\}$, $\Pi_K := \{(\mathtt{pi}_a, \mathtt{id}_{\mathtt{Im}(\mathtt{pi}_a)}) \mid a \in \mathcal{A}\}$, *and* $\Pi_B := \emptyset$; *and an interpretation function* ι *on types and typed function names such that* $\iota(\mathtt{Im}(\mathtt{pi}_a)) := \mathrm{Im}(\pi_a)$ *and* $\iota(\mathtt{pi}_a) := \pi_a$ *, respectively.*

The resulting instantiation structure is $(\mathcal{S}, \iota)$.

2. Epistemic Logic as Dynamic Logic [Par91]: *Recall Parikh's embedding* θ *[Par91] of Epistemic Logic [FHMV95] into Propositional Dynamic Logic [HKT00] with inverse actions [Par78], by which Parikh established an upper, EXPTIME complexity bound for Epistemic Logic (also with common knowledge):*

$$\begin{aligned} \theta(P) &:= P \\ \theta(\neg\phi) &:= \neg\theta(\phi) \\ \theta(\phi \wedge \phi') &:= \theta(\phi) \wedge \theta(\phi') \\ \theta(\mathsf{K}_a(\phi)) &:= [(\alpha_a \cup (\alpha_a)^{-1})^*]\theta(\phi)\,, \end{aligned}$$

where $[(\alpha_a \cup (\alpha_a)^{-1})^*]$ *is the dynamic necessity modality with the program parameter* $(\alpha_a \cup (\alpha_a)^{-1})^*$ *for* α_a *as before. Further, let* α *denote actions, and* A *and* A' *action terms such as* $(\alpha_a \cup (\alpha_a)^{-1})^*$, *and let*

$$\begin{aligned} \mathcal{S} \ni s &::= \mathsf{0} \mid \alpha_a(s) & \mathrm{R}_{A^0} &:= \mathrm{id}_{\mathcal{S}} \\ \mathrm{R}_{\alpha} &:= \{(s, \alpha(s)) \mid s \in \mathcal{S}\} & \mathrm{R}_{A^1} &:= \mathrm{R}_A \\ \mathrm{R}_{A^{-1}} &:= (\mathrm{R}_A)^{-1} & \mathrm{R}_{A^{n+1}} &:= \mathrm{R}_{A^n} \circ \mathrm{R}_{A^1} \\ \mathrm{R}_{A \cup A'} &:= \mathrm{R}_A \cup \mathrm{R}_{A'} & \mathrm{R}_{A^*} &:= \bigcup_{n \in \mathbb{N}} \mathrm{R}_{A^n}\,. \end{aligned}$$

Then,

$$\boxed{\mathrm{id}_{\mathcal{S}} \cup \mathrm{E}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}} = \mathrm{R}_{(\alpha_a \cup (\alpha_a)^{-1})^*}\,,}$$

where $\mathrm{R}_{(\alpha_a \cup (\alpha_a)^{-1})^*}$ *is of course an equivalence relation.*

Notice that R_{α_a} *is not idempotent, and so* $(\mathrm{id}_S, \mathrm{R}_{\alpha_a})$ *is not a doxastic-epistemic function pair, but fortunately thanks to Theorem 2.2, there is a constructible* $\pi_a : \mathcal{S} \to \mathcal{S}$ *such that*

(a) $\mathrm{E}_{\pi_a}^{\mathrm{id}_{\mathcal{S}}} = \mathrm{id}_{\mathcal{S}} \cup \mathrm{E}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}}$, *and*

(b) $(\pi_a, \mathrm{id}_{\mathcal{S}})$ *is a doxastic-epistemic function pair.*

3. *Interactive Provability as Explicit Belief [Kra12b]: From [Kra12b], recall the definition of the (idempotent) operator $\sigma_a^M : \mathcal{S} \to \mathcal{S}$ defined on $\mathcal{S} \ni s ::= \mathsf{0} \mid \mathsf{succ}_a^M(s)$ such that*

$$\sigma_a^M(s) := \begin{cases} s & \text{if } M \in \mathrm{cl}_a^s(\emptyset), \text{ and} \\ \mathsf{succ}_a^M(s) & \text{otherwise (oracle input),} \end{cases}$$

where M designates a proof term, succ_a^M a state constructor, and cl_a^s a closure operator such that $M \in \mathrm{cl}_a^{\mathsf{succ}_a^M(s)}(\emptyset)$. (Here, the exact nature of M and cl_a^s is unimportant.) Then,

$$\boxed{\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\sigma_a^M} = {}_M\mathrm{R}_a\,,}$$

where ${}_M\mathrm{R}_a$ is the accessibility relation for the negation-complete proof modality axiomatised in [Kra12b], obtainable directly as $\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\sigma_a^M}$.

Proof. For 1, inspect Definition 4 and 1.

For 2, consider:

$$\begin{aligned}
\mathrm{id}_{\mathcal{S}} \cup \mathrm{E}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}} &= \mathrm{id}_{\mathcal{S}} \cup (\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}} \cup (\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}})^{-1})^{+} \\
&= (\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}} \cup (\mathrm{D}_{\mathrm{id}_{\mathcal{S}}}^{\mathrm{R}_{\alpha_a}})^{-1})^{*} \\
&= (\mathrm{R}_{\alpha_a} \cup (\mathrm{R}_{\alpha_a})^{-1})^{*} \qquad (\mathrm{R}_{\alpha_a} \text{ is functional}) \\
&= (\mathrm{R}_{\alpha_a} \cup \mathrm{R}_{(\alpha_a)^{-1}})^{*} \\
&= (\mathrm{R}_{\alpha_a \cup (\alpha_a)^{-1}})^{*} \\
&= \mathrm{R}_{(\alpha_a \cup (\alpha_a)^{-1})^{*}}\,.
\end{aligned}$$

For 3, inspect Definition 3 and [Kra12b]. □

3 Conclusion

We conclude by mentioning that from D_f^g and E_f^g, we can further construct accessibility relations for modalities of *common and distributed belief and knowledge* in a standard way [FHMV95, MV07] by taking unions and transitive closures of D_f^g-relations for common belief, unions and reflexive-transitive closures of E_f^g-relations for common knowledge, and intersections of D_f^g- and E_f^g-relations for distributed belief and knowledge, respectively.

For example, accessibility relations $\mathrm{DD}_{\mathcal{C}}^g$ for distributed belief, $\mathrm{CD}_{\mathcal{C}}^g$ for common belief, $\mathrm{DE}_{\mathcal{C}}^g$ for distributed knowledge, and $\mathrm{CE}_{\mathcal{C}}^g$ for common knowledge with respect to doxastic-epistemic function pairs (π_a, g) and a community $\mathcal{C}$ of agents a can be constructed:

$$\begin{array}{lcllcl}
\mathrm{DD}_{\mathcal{C}}^g & := & \bigcap_{a\in\mathcal{C}} \mathrm{D}_{\pi_a}^g & \mathrm{CD}_{\mathcal{C}}^g & := & (\bigcup_{a\in\mathcal{C}} \mathrm{D}_{\pi_a}^g)^{+} \\
\mathrm{DE}_{\mathcal{C}}^g & := & \bigcap_{a\in\mathcal{C}} \mathrm{E}_{\pi_a}^g & \mathrm{CE}_{\mathcal{C}}^g & := & (\bigcup_{a\in\mathcal{C}} \mathrm{E}_{\pi_a}^g)^{*} \\
\mathrm{DE}_{\mathcal{C}} & := & \mathrm{DE}_{\mathcal{C}}^{\mathrm{id}_{\mathcal{S}}} & \mathrm{CE}_{\mathcal{C}} & := & \mathrm{CE}_{\mathcal{C}}^{\mathrm{id}_{\mathcal{S}}}\,.
\end{array}$$

References

[BvB07] P. Blackburn and J. van Benthem. *Handbook of Modal Logic*, chapter Modal Logic: A Semantic Perspective. Volume 3 of Blackburn et al. [BvBW07], 2007.

[BvBW07] P. Blackburn, J. van Benthem, and F. Wolter, editors. *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*. Elsevier, 2007.

[Fef89] S. Feferman. *The Number Systems: Foundations of Algebra and Analysis*. AMS Chelsea Publishing, second edition, 1964 (1989). Reprinted by the American Mathematical Society, 2003.

[FHMV95] R. Fagin, J.Y. Halpern, Y. Moses, and M.Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

[HKT00] D. Harel, D. Kozen, and J. Tiuryn. *Dynamic Logic*. MIT Press, 2000.

[HR10] V.F. Hendricks and O. Roy, editors. *Epistemic Logic: 5 Questions*. Automatic Press, 2010.

[HSS09] J.Y. Halpern, D. Samet, and E. Segev. Defining knowledge in terms of belief: The modal logic perspective. *The Review of Symbolic Logic*, 2(3), 2009.

[Kra12a] S. Kramer. A logic of interactive proofs (formal theory of knowledge transfer). Technical Report 1201.3667, arXiv, 2012. `http://arxiv.org/abs/1201.3667`.

[Kra12b] S. Kramer. Logic of negation-complete interactive proofs (formal theory of epistemic deciders). Technical Report 1208.5913, arXiv, 2012. `http://arxiv.org/abs/1208.5913`.

[Kra12c] S. Kramer. Logic of non-monotonic interactive proofs (formal theory of temporary knowledge transfer). Technical Report 1208.1842, arXiv, 2012. `http://arxiv.org/abs/1208.1842`.

[Mos06] Y. Moschovakis. *Notes on Set Theory*. Springer, 2nd edition, 2006.

[MV07] J.-J. Meyer and F. Veltnam. *Handbook of Modal Logic*, chapter Intelligent Agents and Common Sense Reasoning. Volume 3 of Blackburn et al. [BvBW07], 2007.

[Par78] R. Parikh. The completeness of propositional dynamic logic. volume 64/1978 of *LNCS*. Springer, 1978.

[Par91] R. Parikh. Monotonic and non-monotonic logics of knowledge. *Fundamenta Informaticae*, 15(3–4), 1991.

[SD08] D.A. Simovici and C. Djeraba. *Mathematical Tools for Data Mining: Set Theory, Partial Orders, Combinatorics*. Springer, 2008.